

Contents

Introduction	1
CRoss Industry Standard Process for datamining	1
Doc header 1	1

Introduction

The dataMineR script toolbox aims to be a efficient set of R & knitr scripts, that can be used by experienced and less experience dataminers. The toolbox uses the best of the R community to efficently analayse any arbritraray dataset and make a predictive model on the target variable. The toolbox uses `R version$version.string`, R-studio and knitr(<http://yihui.name/knitr/>) to knit R code and Latex into nice and readable pdf reports. We have the option to include all R code that is used to generate the plots and calculations (see “chunk_options”). By default this feature is dissabled.

CRoss Industry Standard Process for datamining

In this toolkit we will use the CRISP methodology to guide the datamining proces.

Doc header 1

Some text explaining the analysis we are doing

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	37.29	1.878	19.86	8.242e-19
wt	-5.344	0.5591	-9.559	1.294e-10

Table 1: Fitting linear model: mpg ~ wt

	speed	dist
****	Min. : 4.0	Min. : 2
****	1st Qu.:12.0	1st Qu.: 26
****	Median :15.0	Median : 36
****	Mean :15.4	Mean : 43
****	3rd Qu.:19.0	3rd Qu.: 56
****	Max. :25.0	Max. :120

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-17.58	6.758	-2.601	0.01232
speed	3.932	0.4155	9.464	1.49e-12

Table 2: Fitting linear model: $\text{dist} \sim \text{speed}$

This report was generated with **R** (2.15.2) and **pander** (0.3.1) on x86_64-apple-darwin9.8.0 platform.

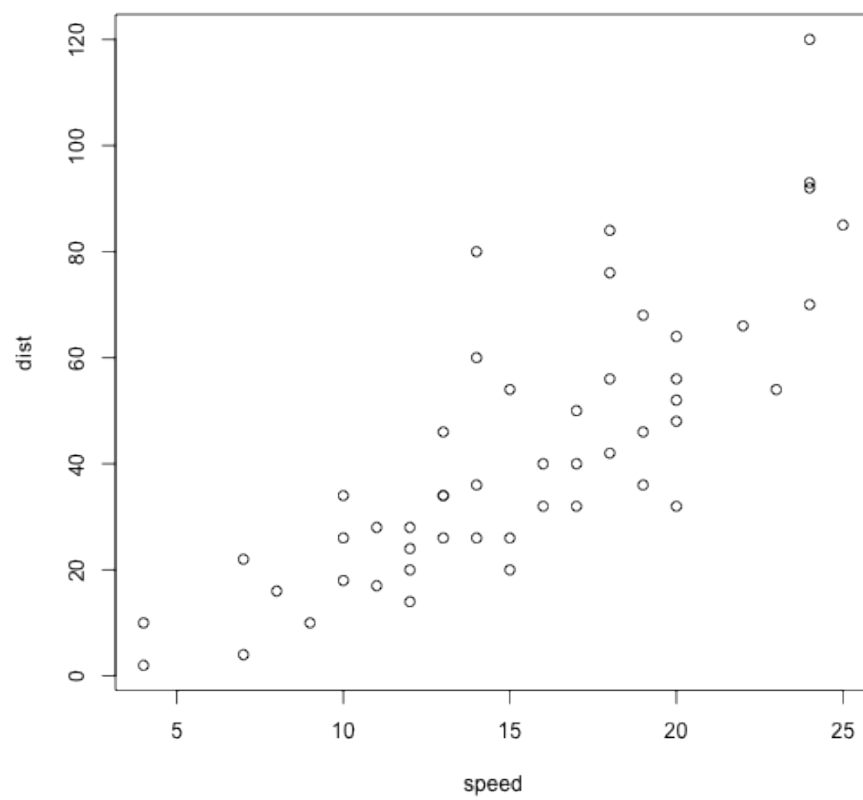


Figure 1: plot of chunk unnamed-chunk-1